

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Ni *et al.*
- (ii) TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR INHIBITOR
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 - (B) STREET: 1100 NEW YORK AVENUE, SUITE 600
 - (C) CITY: WASHINGTON
 - (D) STATE: DC
 - (E) COUNTRY: USA
 - (F) ZIP: 20005-3934
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US To be assigned
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/934,011
 - (B) FILING DATE: 15-AUG-1997
- (viii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/024,056
 - (B) FILING DATE: 16-AUG-1996
- (ix) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: STEFFE, ERIC K.
 - (B) REGISTRATION NUMBER: 36,688
 - (C) REFERENCE/DOCKET NUMBER: 1488.0300002
- (x) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 202-371-2600
 - (B) TELEFAX: 202-371-2540

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1371 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: both
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS

(B) LOCATION: 67..1242

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 109..1242

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 67..108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCACGAGGG AAAACTCTAT TTTGAAAATG AATATATTTT GATTAAACAA ATACAGAGAA	60
GTCAAA ATG GAC ACA ATC TTC TTG TGG AGT CTT CTA TTG CTG TTT TTT Met Asp Thr Ile Phe Leu Trp Ser Leu Leu Leu Phe Phe	108
-14 -10 -5	
GGA AGT CAA GCC TCA AGA TGC TCA GCT CAA AAA AAT ACC GAA TTT GCA Gly Ser Gln Ala Ser Arg Cys Ser Ala Gln Lys Asn Thr Glu Phe Ala	156
1 5 10 15	
GTG GAT CTT TAT CAA GAG GTT TCC TTA TCT CAT AAG GAC AAC ATT ATA Val Asp Leu Tyr Gln Glu Val Ser Leu Ser His Lys Asp Asn Ile Ile	204
20 25 30	
TTT TCA CCC CTT GGA ATA ACT TTG GTT CTT GAG ATG GTA CAA CTG GGA Phe Ser Pro Leu Gly Ile Thr Leu Val Leu Glu Met Val Gln Leu Gly	252
35 40 45	
GCC AAA GGA AAA GCA CAG CAG CAG ATA AGA CAA ACT TTA AAA CAA CAG Ala Lys Gly Lys Ala Gln Gln Ile Arg Gln Thr Leu Lys Gln Gln	300
50 55 60	
GAA ACC TCA GCT GGG GAA GAA TTT TTG GTA CTG AAG TCA TTT TGC TCT Glu Thr Ser Ala Gly Glu Glu Phe Leu Val Leu Lys Ser Phe Cys Ser	348
65 70 75 80	
GCC ATC TCA GAG AAA AAA CAA GAA TTT ACA TTT AAT CTT GCC AAT GCC Ala Ile Ser Glu Lys Gln Glu Phe Thr Phe Asn Leu Ala Asn Ala	396
85 90 95	
CTC TAC CTT CAA GAA GGA TTC ACT GTG AAA GAA CAG TAT CTC CAT GGC Leu Tyr Leu Gln Glu Gly Phe Thr Val Lys Glu Gln Tyr Leu His Gly	444
100 105 110	
AAC AAG GAA TTT TTT CAG AGT GCT ATA AAA CTG GTG GAT TTT CAA GAT Asn Lys Glu Phe Phe Gln Ser Ala Ile Lys Leu Val Asp Phe Gln Asp	492
115 120 125	
GCA AAG GCT TGT GCA GAG ATG ATA AGT ACC TGG GTA GAA AGA AAA ACA Ala Lys Ala Cys Ala Glu Met Ile Ser Thr Trp Val Glu Arg Lys Thr	540
130 135 140	
GAT GGA AAA ATT AAA GAC ATG TTT TCA GGG GAA GAA TTT GGC CCT CTG Asp Gly Lys Ile Lys Asp Met Phe Ser Gly Glu Glu Phe Gly Pro Leu	588
145 150 155 160	
ACT CGG CTT GTC CTG GTG AAT GCT ATT TAT TTC AAA GGA GAT TGG AAA Thr Arg Leu Val Leu Val Asn Ala Ile Tyr Phe Lys Gly Asp Trp Lys	636
165 170 175	

CAG AAA TTC AGA AAA GAG GAC ACA CAG CTG ATA AAT TTT ACT AAG AAA Gln Lys Phe Arg Lys Glu Asp Thr Gln Leu Ile Asn Phe Thr Lys Lys 180 185 190	684
AAT GGT TCA ACT GTC AAA ATT CCA ATG ATG AAG GCT CTT CTG AGA ACA Asn Gly Ser Thr Val Lys Ile Pro Met Met Lys Ala Leu Leu Arg Thr 195 200 205	732
AAA TAT GGT TAT TTT TCT GAA TCT TCC CTG AAC TAC CAA GTT TTA GAA Lys Tyr Gly Tyr Phe Ser Glu Ser Ser Leu Asn Tyr Gln Val Leu Glu 210 215 220	780
TTG TCT TAC AAA GGT GAT GAA TTT AGC TTA ATT ATC ATA CTT CCT GCA Leu Ser Tyr Lys Gly Asp Glu Phe Ser Leu Ile Ile Ile Leu Pro Ala 225 230 235 240	828
GAA GGT ATG GAT ATA GAA GAA GTG GAA AAA CTA ATT ACT GCT CAA CAA Glu Gly Met Asp Ile Glu Glu Val Glu Lys Leu Ile Thr Ala Gln Gln 245 250 255	876
ATC CTA AAA TGG CTC TCT GAG ATG CAA GAA GAG GAA GTA GAA ATA AGC Ile Leu Lys Trp Leu Ser Glu Met Gln Glu Glu Glu Val Glu Ile Ser 260 265 270	924
CTC CCT AGA TTT AAA GTA GAA CAA AAA GTA GAC TTC AAA GAC GTT TTG Leu Pro Arg Phe Lys Val Glu Gln Lys Val Asp Phe Lys Asp Val Leu 275 280 285	972
TAT TCT TTG AAC ATA ACC GAG ATA TTT AGT GGT GGC TGC GAC CTT TCT Tyr Ser Leu Asn Ile Thr Glu Ile Phe Ser Gly Gly Cys Asp Leu Ser 290 295 300	1020
GGA ATA ACA GAT TCA TCT GAA GTG TAT GTT TCC CAA GTG ACG CAA AAA Gly Ile Thr Asp Ser Ser Glu Val Tyr Val Ser Gln Val Thr Gln Lys 305 310 315 320	1068
GTT TTC TTT GAG ATA AAT GAA GAT GGT AGT GAA GCT GCA ACA TCA ACT Val Phe Phe Glu Ile Asn Glu Asp Gly Ser Glu Ala Ala Thr Ser Thr 325 330 335	1116
GGC ATA CAC ATC CCT GTG ATC ATG AGT CTG GCT CAA AGC CAA TTT ATA Gly Ile His Ile Pro Val Ile Met Ser Leu Ala Gln Ser Gln Phe Ile 340 345 350	1164
GCA AAT CAT CCA TTT CTG TTT ATT ATG AAG CAT AAT CCA ACA GAA TCA Ala Asn His Pro Phe Leu Phe Ile Met Lys His Asn Pro Thr Glu Ser 355 360 365	1212
ATT CTG TTT ATG GGA AGA GTG ACA AAT CCC TGACACCCAG GAGATAAAAG Ile Leu Phe Met Gly Arg Val Thr Asn Pro 370 375	1262
GAAGAGAGTTT AGATTCACTG TGAATGAAAAA GCACAGCCTC AGAATAAAAG ATGATTTCTC	1322
AAAAATAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1371

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 392 amino acids
 - (B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Thr Ile Phe Leu Trp Ser Leu Leu Leu Phe Phe Gly Ser
-14 -10 -5 1

Gln Ala Ser Arg Cys Ser Ala Gln Lys Asn Thr Glu Phe Ala Val Asp
5 10 15

Leu Tyr Gln Glu Val Ser Leu Ser His Lys Asp Asn Ile Ile Phe Ser
20 25 30

Pro Leu Gly Ile Thr Leu Val Leu Glu Met Val Gln Leu Gly Ala Lys
35 40 45 50

Gly Lys Ala Gln Gln Ile Arg Gln Thr Leu Lys Gln Gln Glu Thr
55 60 65

Ser Ala Gly Glu Glu Phe Leu Val Leu Lys Ser Phe Cys Ser Ala Ile
70 75 80

Ser Glu Lys Lys Gln Glu Phe Thr Phe Asn Leu Ala Asn Ala Leu Tyr
85 90 95

Leu Gln Glu Gly Phe Thr Val Lys Glu Gln Tyr Leu His Gly Asn Lys
100 105 110

Glu Phe Phe Gln Ser Ala Ile Lys Leu Val Asp Phe Gln Asp Ala Lys
115 120 125 130

Ala Cys Ala Glu Met Ile Ser Thr Trp Val Glu Arg Lys Thr Asp Gly
135 140 145

Lys Ile Lys Asp Met Phe Ser Gly Glu Phe Gly Pro Leu Thr Arg
150 155 160

Leu Val Leu Val Asn Ala Ile Tyr Phe Lys Gly Asp Trp Lys Gln Lys
165 170 175

Phe Arg Lys Glu Asp Thr Gln Leu Ile Asn Phe Thr Lys Lys Asn Gly
180 185 190

Ser Thr Val Lys Ile Pro Met Met Lys Ala Leu Leu Arg Thr Lys Tyr
195 200 205 210

Gly Tyr Phe Ser Glu Ser Ser Leu Asn Tyr Gln Val Leu Glu Leu Ser
215 220 225

Tyr Lys Gly Asp Glu Phe Ser Leu Ile Ile Leu Pro Ala Glu Gly
230 235 240

Met Asp Ile Glu Glu Val Glu Lys Leu Ile Thr Ala Gln Gln Ile Leu
245 250 255

Lys Trp Leu Ser Glu Met Gln Glu Glu Glu Val Glu Ile Ser Leu Pro
260 265 270

Arg Phe Lys Val Glu Gln Lys Val Asp Phe Lys Asp Val Leu Tyr Ser
275 280 285 290

Leu Asn Ile Thr Glu Ile Phe Ser Gly Gly Cys Asp Leu Ser Gly Ile
 295 300 305
 Thr Asp Ser Ser Glu Val Tyr Val Ser Gln Val Thr Gln Lys Val Phe
 310 315 320
 Phe Glu Ile Asn Glu Asp Gly Ser Glu Ala Ala Thr Ser Thr Gly Ile
 325 330 335
 His Ile Pro Val Ile Met Ser Leu Ala Gln Ser Gln Phe Ile Ala Asn
 340 345 350
 His Pro Phe Leu Phe Ile Met Lys His Asn Pro Thr Glu Ser Ile Leu
 355 360 365 370
 Phe Met Gly Arg Val Thr Asn Pro
 375

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 402 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gln Met Ser Pro Ala Leu Thr Cys Leu Val Leu Gly Leu Ala Leu
 1 5 10 15
 Val Phe Gly Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala
 20 25 30
 His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gln Gln Val Ala Gln
 35 40 45
 Ala Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser
 50 55 60
 Val Leu Ala Met Leu Gln Leu Thr Thr Gly Gly Glu Thr Gln Gln Gln
 65 70 75 80
 Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro
 85 90 95
 Ala Leu Arg His Leu Tyr Lys Glu Leu Met Gly Pro Trp Asn Lys Asp
 100 105 110
 Glu Ile Ser Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu
 115 120 125
 Val Gln Gly Phe Met Pro His Phe Phe Arg Leu Phe Arg Ser Thr Val
 130 135 140
 Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile Asn

145	150	155	160
Asp Trp Val Lys Thr His Thr Lys Gly Met Ile Ser Asn Leu Leu Gly			
165		170	175
Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala Leu			
180	185		190
Tyr Phe Asn Gly Gln Trp Lys Thr Pro Phe Pro Asp Ser Ser Thr His			
195	200		205
Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser Val Pro Met			
210	215		220
Met Ala Gln Thr Asn Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp			
225	230	235	240
Gly His Tyr Tyr Asp Ile Leu Glu Leu Pro Tyr His Gly Asp Thr Leu			
245		250	255
Ser Met Phe Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser Ala			
260		265	270
Leu Thr Asn Ile Leu Ser Ala Gln Leu Ile Ser His Trp Lys Gly Asn			
275	280		285
Met Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu			
290	295		300
Thr Glu Val Asp Leu Arg Lys Pro Leu Glu Asn Leu Gly Met Thr Asp			
305	310	315	320
Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu Ser Asp Gln Glu			
325		330	335
Pro Leu His Val Ala Gln Ala Leu Gln Lys Val Lys Ile Glu Val Asn			
340		345	350
Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile Val Ser Ala			
355	360		365
Arg Met Ala Pro Glu Glu Ile Ile Met Asp Arg Pro Phe Leu Phe Val			
370	375		380
Val Arg His Asn Pro Thr Gly Thr Val Leu Phe Met Gly Gln Val Met			
385	390	395	400
Glu Pro			

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 415 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Asp Leu Cys Val Ala Asn Thr Leu Phe Ala Leu Asn Leu Phe
1 5 10 15

Lys His Leu Ala Lys Ala Ser Pro Thr Gln Asn Leu Phe Leu Ser Pro
20 25 30

Trp Ser Ile Ser Ser Thr Met Ala Met Val Tyr Met Gly Ser Arg Gly
35 40 45

Ser Thr Glu Asp Gln Met Ala Lys Val Leu Gln Phe Asn Glu Val Gly
50 55 60

Ala Asn Ala Val Thr Pro Met Thr Pro Glu Asn Phe Thr Ser Cys Gly
65 70 75 80

Phe Met Gln Gln Ile Gln Lys Gly Ser Tyr Pro Asp Ala Ile Leu Gln
85 90 95

Ala Gln Ala Ala Asp Lys Ile His Ser Ser Phe Arg Ser Leu Ser Ser
100 105 110

Ala Ile Asn Ala Ser Thr Gly Asp Tyr Leu Leu Glu Ser Val Asn Lys
115 120 125

Leu Phe Gly Glu Lys Ser Ala Ser Phe Arg Glu Glu Tyr Ile Arg Leu
130 135 140

Cys Gln Lys Tyr Tyr Ser Ser Glu Pro Gln Ala Val Asp Phe Leu Glu
145 150 155 160

Cys Ala Glu Glu Ala Arg Lys Lys Ile Asn Ser Trp Val Lys Thr Gln
165 170 175

Thr Lys Gly Lys Ile Pro Asn Leu Leu Pro Glu Gly Ser Val Asp Gly
180 185 190

Asp Thr Arg Met Val Leu Val Asn Ala Val Tyr Phe Lys Gly Lys Trp
195 200 205

Lys Thr Pro Phe Glu Lys Lys Leu Asn Gly Leu Tyr Pro Phe Arg Val
210 215 220

Asn Ser Ala Gln Arg Thr Pro Val Gln Met Met Tyr Leu Arg Glu Lys
225 230 235 240

Leu Asn Ile Gly Tyr Ile Glu Asp Leu Lys Ala Gln Ile Leu Glu Leu
245 250 255

Pro Tyr Ala Gly Asp Val Ser Met Phe Leu Leu Leu Pro Asp Glu Ile
260 265 270

Ala Asp Val Ser Thr Gly Leu Glu Leu Leu Glu Ser Glu Ile Thr Tyr
275 280 285

Asp Lys Leu Asn Lys Trp Thr Ser Lys Asp Lys Met Ala Glu Asp Glu
290 295 300

Val Glu Val Tyr Ile Pro Gln Phe Lys Leu Glu Glu His Tyr Glu Leu
305 310 315 320

Arg Ser Ile Leu Arg Ser Met Gly Met Glu Asp Ala Phe Asn Lys Gly
 325 330 335
 Arg Ala Asn Phe Ser Gly Met Ser Glu Arg Asn Asp Leu Phe Leu Ser
 340 345 350
 Glu Val Phe His Gln Ala Met Val Asp Val Asn Glu Glu Gly Thr Glu
 355 360 365
 Ala Ala Ala Gly Thr Gly Gly Val Met Thr Gly Arg Thr Gly His Gly
 370 375 380
 Gly Pro Gln Phe Val Ala Asp His Pro Phe Leu Phe Leu Ile Met His
 385 390 395 400
 Lys Ile Thr Lys Cys Ile Leu Phe Phe Gly Arg Phe Cys Ser Pro
 405 410 415

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCCCCATGGG AAGTCAAGCC TCAAG 25

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCAAGCTTT CACTTCCTTT TATCTCCCTG 30

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGATCCG CCATCATGGA CACAATCTTC TTG

33

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGGTACCT CACTCCTTT TATCTCCCTG

30

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 57 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCTCTAGAT CAAGCGTAGT CTGGGACGTC GTATGGGTAG GGATTGTCA CTCTTCC

57

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 171 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

NAATATATTT NNATTTAAC AATAACAGAGA AGTCAAAATG GACACAATCT TCTTGTTGGAG

60

TCTTCTATTG CTGTTTTTC GAAGTCAAGC CTCANGAATG CTCAGCTGCA AAAAATACC

120

GAATTTGCCA GTGGNATCTT TATCAAGAGG TTTCCCTTCAT CTGCATAAGG N

171

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCANNANAA CAATCTNATC CAAGGACTGT GGNACTCCTG TTCCCTGCTC ATCATGTCAT	60
GGGGCATCTG CCAGGAACCA TCTTGATGG TGTA _{AAA} ATC TTGAATACAT AAGAGGGAAA	120
TTTAGACTT GTTAGAAAGA AGCCAAGCAA TTGAGACCTT AGATAGAACT TAGAATTCTC	180
GCCGAGTTT GTTGGTAAT TGTTACTTCA AAA _{AAA} ATG CAATTCTGT TCCCTCTTC	240
CTCCAACCAT TTATCTGGGA AGCAAGTTAT TGGCAACCCA GAGCTGATTG TTGGAGCCGG	300
GGAAAATGGT GTGAAATGTG AGAAAATGTA ATTGAGATAA TAAAAACAAA AGATTTACA	360
ATATATTATC CTCTAAGTCA TCCATTA _{AAA} A ATTGGTAGC AAA _{AA} TGTGC AGTGT _{TT} CAA	420
GACTTTCTT TTCTTTTTT TTNAATACCA GATTAAAGTA GACCA _{AA} AG TAGACTCCAA	480
AGACGTTTGG ATNCTTGAAC ATAACCGNGA TATTA	515

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1370 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 67..1281

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
(B) LOCATION: 67..120

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
(B) LOCATION: 121..1281

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCACGAGGG AAAACTCTAT TTTGAAAATG AATATATT _{TTT} GATTAAACA ATACAGAGAA	60
GTCAAA ATG GAC ACA ATC TTC TTG TGG AGT CTT CTA TTG CTG TTT TTT	108

Met Asp Thr Ile Phe Leu Trp Ser Leu Leu Leu Leu Phe Phe	-18	-15	-10	-5	
Gly Ser Gln Ala Ser Arg Cys Ser Ala Gln Lys Asn Thr Glu Phe Ala	1	5	10		156
GTG GAT CTT TAT CAA GAG GTT TCC TTA TCT CAT AAG GAC AAC ATT ATA	15	20	25		204
Val Asp Leu Tyr Gln Glu Val Ser Leu Ser His Lys Asp Asn Ile Ile					
TTT TCA CCC CTT GGA ATA ACT TTG GTT CTT GAG ATG GTA CAA CTG GGA	30	35	40		252
Phe Ser Pro Leu Gly Ile Thr Leu Val Leu Glu Met Val Gln Leu Gly					
GCC AAA GGA AAA GCA CAG CAG ATA AGA CAA ACT TTA AAA CAA CAG	45	50	55	60	300
Ala Lys Gly Lys Ala Gln Gln Ile Arg Gln Thr Leu Lys Gln Gln					
GAA ACC TCA GCT GGG GAA GAA TTT TTG GTA CTG AAG TCA TTT TGC TCT	65	70	75		348
Glu Thr Ser Ala Gly Glu Glu Phe Leu Val Leu Lys Ser Phe Cys Ser					
GCC ATC TCA GAG AAA AAA CAA GAA TTT ACA TTT AAT CTT GCC AAT GCC	80	85	90		396
Ala Ile Ser Glu Lys Gln Glu Phe Thr Phe Asn Leu Ala Asn Ala					
CTC TAC CTT CAA GAA GGA TTC ACT GTG AAA GAA CAG TAT CTC CAT GGC	95	100	105		444
Leu Tyr Leu Gln Glu Gly Phe Thr Val Lys Glu Gln Tyr Leu His Gly					
AAC AAG GAA TTT TTT CAG AGT GCT ATA AAA CTG GTG GAT TTT CAA GAT	110	115	120		492
Asn Lys Glu Phe Phe Gln Ser Ala Ile Lys Leu Val Asp Phe Gln Asp					
GCA AAG GCT TGT GCA GAG ATG ATA AGT ACC TGG GTA GAA AGA AAA ACA	125	130	135	140	540
Ala Lys Ala Cys Ala Glu Met Ile Ser Thr Trp Val Glu Arg Lys Thr					
GAT GGA AAA ATT AAA GAC ATG TTT TCA GGG GAA GAA TTT GGC CCT CTG	145	150	155		588
Asp Gly Lys Ile Lys Asp Met Phe Ser Gly Glu Glu Phe Gly Pro Leu					
ACT CGG CTT GTC CTG GTG AAT GCT ATT TAT TTC AAA GGA GAT TGG AAA	160	165	170		636
Thr Arg Leu Val Leu Val Asn Ala Ile Tyr Phe Lys Gly Asp Trp Lys					
CAG AAA TTC AGA AAA GAG GAC ACA CAG CTG ATA AAT TTT ACT AAG AAA	175	180	185		684
Gln Lys Phe Arg Lys Glu Asp Thr Gln Leu Ile Asn Phe Thr Lys Lys					
AAT GGT TCA ACT GTC AAA ATT CCA ATG ATG AAG GCT CTT CTG AGA ACA	190	195	200		732
Asn Gly Ser Thr Val Lys Ile Pro Met Met Lys Ala Leu Leu Arg Thr					
AAA TAT GGT TAT TTT TCT GAA TCT TCC CTG AAC TAC CAA GTT TTA GAA	205	210	215	220	780
Lys Tyr Gly Tyr Phe Ser Glu Ser Ser Leu Asn Tyr Gln Val Leu Glu					
TTG TCT TAC AAA GGT GAT GAA TTT AGC TTA ATT ATC ATA CTT CCT GCA	225	230	235		828
Leu Ser Tyr Lys Gly Asp Glu Phe Ser Leu Ile Ile Ile Leu Pro Ala					

GAA GGT ATG GAT ATA GAA GAA GTG GAA AAA CTA ATT ACT GCT CAA CAA Glu Gly Met Asp Ile Glu Glu Val Glu Lys Leu Ile Thr Ala Gln Gln 240 245 250	876
ATC CTA AAA TGG CTC TCT GAG ATG CAA GAA GAG GAA GTA GAA ATA AGC Ile Leu Lys Trp Leu Ser Glu Met Gln Glu Glu Glu Val Glu Ile Ser 255 260 265	924
CTC CCT AGA TTT AAA GTA GAA CAA AAA GTA GAC TTC AAA GAC GTT TTG Leu Pro Arg Phe Lys Val Glu Gln Lys Val Asp Phe Lys Asp Val Leu 270 275 280	972
TAT TCT TTG AAC ATA ACC GAG ATA TTT AGT GGT GGC TGC GAC CTT TCT Tyr Ser Leu Asn Ile Thr Glu Ile Phe Ser Gly Gly Cys Asp Leu Ser 285 290 295 300	1020
GGA ATA ACA GAT TCA TCT GAA GTG TAT GTT TCC CAA GTG ACG CAA AAA Gly Ile Thr Asp Ser Ser Glu Val Tyr Val Ser Gln Val Thr Gln Lys 305 310 315	1068
GTT TTC TTT GAG ATA AAT GAA GAT GGT AGT GAA GCT GCA ACA TCA ACT Val Phe Glu Ile Asn Glu Asp Gly Ser Glu Ala Ala Thr Ser Thr 320 325 330	1116
GGC ATA CAC ATC CCT GTG ATC ATG AGT CTG GCT CAA AGC CAA TTT ATA Gly Ile His Ile Pro Val Ile Met Ser Leu Ala Gln Ser Gln Phe Ile 335 340 345	1164
GCA AAT CAT CCA TTT CTG TTT ATT ATG AAG CAT AAT CCA ACA GAA TCA Ala Asn His Pro Phe Leu Phe Ile Met Lys His Asn Pro Thr Glu Ser 350 355 360	1212
ATT CTG TTT ATG GGA AGA GTG ACA AAT CCT GAC ACC CAG GAG ATA AAA Ile Leu Phe Met Gly Arg Val Thr Asn Pro Asp Thr Gln Glu Ile Lys 365 370 375 380	1260
GGA AGA GAT TTA GAT TCA CTG TGAATGAAAAA GCACAGCCTC AGAATAAAAG Gly Arg Asp Leu Asp Ser Leu 385	1311
ATGATTTCTC AAAAATAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1370

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 405 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Asp Thr Ile Phe Leu Trp Ser Leu Leu Leu Phe Phe Gly Ser -18 -15 -10 -5
Gln Ala Ser Arg Cys Ser Ala Gln Lys Asn Thr Glu Phe Ala Val Asp 1 5 10
Leu Tyr Gln Glu Val Ser Leu Ser His Lys Asp Asn Ile Ile Phe Ser 15 20 25 30

Pro Leu Gly Ile Thr Leu Val Leu Glu Met Val Gln Leu Gly Ala Lys
 35 40 45
 Gly Lys Ala Gln Gln Gln Ile Arg Gln Thr Leu Lys Gln Gln Glu Thr
 50 55 60
 Ser Ala Gly Glu Glu Phe Leu Val Leu Lys Ser Phe Cys Ser Ala Ile
 65 70 75
 Ser Glu Lys Lys Gln Glu Phe Thr Phe Asn Leu Ala Asn Ala Leu Tyr
 80 85 90
 Leu Gln Glu Gly Phe Thr Val Lys Glu Gln Tyr Leu His Gly Asn Lys
 95 100 105 110
 Glu Phe Phe Gln Ser Ala Ile Lys Leu Val Asp Phe Gln Asp Ala Lys
 115 120 125
 Ala Cys Ala Glu Met Ile Ser Thr Trp Val Glu Arg Lys Thr Asp Gly
 130 135 140
 Lys Ile Lys Asp Met Phe Ser Gly Glu Glu Phe Gly Pro Leu Thr Arg
 145 150 155
 Leu Val Leu Val Asn Ala Ile Tyr Phe Lys Gly Asp Trp Lys Gln Lys
 160 165 170
 Phe Arg Lys Glu Asp Thr Gln Leu Ile Asn Phe Thr Lys Lys Asn Gly
 175 180 185 190
 Ser Thr Val Lys Ile Pro Met Met Lys Ala Leu Leu Arg Thr Lys Tyr
 195 200 205
 Gly Tyr Phe Ser Glu Ser Ser Leu Asn Tyr Gln Val Leu Glu Leu Ser
 210 215 220
 Tyr Lys Gly Asp Glu Phe Ser Leu Ile Ile Leu Pro Ala Glu Gly
 225 230 235
 Met Asp Ile Glu Glu Val Glu Lys Leu Ile Thr Ala Gln Gln Ile Leu
 240 245 250
 Lys Trp Leu Ser Glu Met Gln Glu Glu Glu Val Glu Ile Ser Leu Pro
 255 260 265 270
 Arg Phe Lys Val Glu Gln Lys Val Asp Phe Lys Asp Val Leu Tyr Ser
 275 280 285
 Leu Asn Ile Thr Glu Ile Phe Ser Gly Gly Cys Asp Leu Ser Gly Ile
 290 295 300
 Thr Asp Ser Ser Glu Val Tyr Val Ser Gln Val Thr Gln Lys Val Phe
 305 310 315
 Phe Glu Ile Asn Glu Asp Gly Ser Glu Ala Ala Thr Ser Thr Gly Ile
 320 325 330
 His Ile Pro Val Ile Met Ser Leu Ala Gln Ser Gln Phe Ile Ala Asn
 335 340 345 350
 His Pro Phe Leu Phe Ile Met Lys His Asn Pro Thr Glu Ser Ile Leu
 355 360 365

Phe Met Gly Arg Val Thr Asn Pro Asp Thr Gln Glu Ile Lys Gly Arg
370 375 380

Asp Leu Asp Ser Leu
385

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3974 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGTACCTAAG	TGAGTAGGGC	GTCCGATCGA	CGGACGCCTT	TTTTTGAAAT	TCGTAATCAT	60
GGTCATAGCT	GTTCCCTGTG	TGAAATTGTT	ATCCGCTCAC	AATTCCACAC	AACATACGAG	120
CCGGAAGCAT	AAAGTGTAAA	GCCTGGGGTG	CCTAATGAGT	GAGCTAACTC	ACATTAATTG	180
CGTTGCGCTC	ACTGCCCGCT	TTCCAGTCGG	GAAACCTGTC	GTGCCAGCTG	CATTAATGAA	240
TCGGCCAACG	CGCGGGGAGA	GGCGGTTTGC	GTATTGGGCG	CTCTTCCGCT	TCCTCGCTCA	300
CTGACTCGCT	GCGCTCGGTC	GTTCGGCTGC	GGCGAGCGGT	ATCAGCTCAC	TCAAAGCGG	360
TAATACGGTT	ATCCACAGAA	TCAGGGATA	ACGCAGGAAA	GAACATGTGA	GCAAAAGGCC	420
AGCAAAAGGC	CAGGAACCGT	AAAAAGGCCG	CGTTGCTGGC	GTTTTCCAT	AGGCTCCGCC	480
CCCCCTGACGA	GCATCACAAA	AATCGACGCT	CAAGTCAGAG	GTGGCGAAC	CCGACAGGAC	540
TATAAAGATA	CCAGGGCGTT	CCCCCTGGAA	GCTCCCTCGT	GCGCTCTCCT	GTTCGACCC	600
TGCCGCTTAC	CGGATACCTG	TCCGCCTTTC	TCCCTCGGG	AAGCGTGGCG	CTTTCTCATA	660
GCTCACGCTG	TAGGTATCTC	AGTTCGGTGT	AGGTCGTTCG	CTCCAAGCTG	GGCTGTGTGC	720
ACGAACCCCC	CGTTCAGCCC	GACCGCTGCG	CCTTATCCGG	TAACTATCGT	CTTGAGTCCA	780
ACCCGGTAAG	ACACGACTTA	TCGCCACTGG	CAGCAGCCAC	TGGTAACAGG	ATTAGCAGAG	840
CGAGGTATGT	AGGCGGTGCT	ACAGAGTTCT	TGAAGTGGTG	GCCTAACTAC	GGCTACACTA	900
GAAGAACAGT	ATTGGTATC	TGCGCTCTGC	TGAAGCCAGT	TACCTTCGGA	AAAAGAGTTG	960
GTAGCTCTTG	ATCCGGCAAA	CAAACCACCG	CTGGTAGCGG	TGGTTTTTTT	GTTCGCAAGC	1020
AGCAGATTAC	GCGCAGAAAA	AAAGGATCTC	AAGAAGATCC	TTTGATCTTT	TCTACGGGT	1080
CTGACGCTCA	GTGGAACGAA	AACTCACGTT	AAGGGATTTT	GGTCATGAGA	TTATCGTCGA	1140
CAATT CGCG	GCGAAGGCCGA	AGCGGCATGC	ATT TACGTTG	ACACC ATCGA	ATGGTGCAAA	1200
ACCTTCGCG	GTATGGCATG	ATAGGCCCG	GAAGAGAGTC	AATT CAGGGT	GGTGAATGTG	1260

AAACCAGTAA CGTTATACGA	TGTCGCAGAG TATGCCGGTG	TCTCTTATCA GACC GTTCC	1320
CGCGTGGTGA ACCAGGCCAG	CCACGTTCT GCGAAAACGC	GGGAAAAAGT GGAAGCGGCG	1380
ATGGCGGAGC TGAATTACAT	TCCCACCGC GTGGCACAAAC	AACTGGCGGG CAAACAGTCG	1440
TTGCTGATTG GCGTTGCCAC	CTCCAGTCTG GCCCTGCACG	CGCCGTGCGA AATTGTGCG	1500
GCGATTAAAT CTCGCGCCGA	TCAACTGGGT CCCAGCGTGG	TGGTGTGAT GGTAGAACGA	1560
AGCGGCGTCG AAGCCTGTAA	AGCGGCGGTG CACAATCTC	TCGCGCAACG CGTCAGTGGG	1620
CTGATCATTA ACTATCCGCT	GGATGACCGAG GATGCCATTG	CTGTGGAAGC TGCGTGCACT	1680
AATGTTCCGG CGTTATTCT	TGATGTCTCT GACCAGACAC	CCATCAACAG TATTATTTTC	1740
TCCCATGAAG ACGGTACGCG	ACTGGCGTG GAGCATCTGG	TCGCATTGGG TCACCAGCAA	1800
ATCGCGCTGT TAGCGGGCCC	ATTAAGTTCT GTCTCGGCC	GTCTGCGTCT GGCTGGCTGG	1860
CATAAAATATC TCACTCGCAA	TCAAATTCAAG CCGATAGCGG	AACGGGAAGG CGACTGGAGT	1920
GCCATGTCCG GTTTCAACA	AACCATGCAA ATGCTGAATG	AGGGCATCGT TCCCAC TGCG	1980
ATGCTGGTTG CCAACGATCA	GATGGCGCTG GGCGCAATGC	GCGCCATTAC CGAGTCCGGG	2040
CTGCGCGTTG GTGCGGATAT	CTCGGTAGTG GGATACGACG	ATACCGAAGA CAGCTCATGT	2100
TATATCCCGC CGTTAACAC	CATCAAACAG GATTTCGCC	TGCTGGGGCA AACCAGCGTG	2160
GACCGCTTGC TGCAACTCTC	TCAGGGCCAG GCGGTGAAGG	GCAATCAGCT GTTGCCCCGTC	2220
TCACTGGTGA AAAGAAAAAC	CACCCTGGCG CCCAATACGC	AAACCGCCTC TCCCCGGCG	2280
TTGGCCGATT CATTAATGCA	GCTGGCACGA CAGGTTCCC	GAUTGGAAAG CGGGCAGTGA	2340
GCGCAACGCA ATTAATGTA	GTTAGCGCGA ATTGTCGACC	AAAGCGGCCA TCGTGCCTCC	2400
CCACTCCTGC AGTCGGGGG	CATGGATGCG CGGATAGCCG	CTGCTGGTT CCTGGATGCC	2460
GACGGATTG CACTGCCGGT	AGAACTCCGC GAGGTCGTCC	AGCCTCAGGC AGCAGCTGAA	2520
CCAACTCGCG AGGGGATCGA	GCCCCGGGTG GGCGAAGAAC	TCCAGCATGA GATCCCCGCG	2580
CTGGAGGATC ATCCAGCCGG	CGTCCCGGAA AACGATTCCG	AAGCCCAACC TTTCATAGAA	2640
GGCGCGGTG GAATCGAAAT	CTCGTGATGG CAGGTTGGC	GTCGCTTGGT CGGTCAATTTC	2700
GAACCCCAGA GTCCCGCTCA	GAAGAACTCG TCAAGAAGGC	GATAGAAGGC GATGCGCTGC	2760
GAATCGGGAG CGCGATACC	GTAAAGCACG AGGAAGCGGT	CAGCCCATTC GCCGCCAACG	2820
TCTTCAGCAA TATCACGGGT	AGCCAACGCT ATGTCCTGAT	AGCGGTCCGC CACACCCAGC	2880
CGGCCACAGT CGATGAATCC	AGAAAAGCGG CCATTTCCA	CCATGATATT CGGCAAGCAG	2940
GCATCGCCAT GGGTCACGAC	GAGATCCTCG CCGTCGGGCA	TGCGCGCCTT GAGCCTGGCG	3000
AACAGTTCGG CTGGCGCGAG	CCCCTGATGC TCTTCGTCCA	GATCATCCTG ATCGACAAGA	3060
CCGGCTTCCA TCCGAGTACG	TGCTCGCTCG ATGCGATGTT	TCGCTTGGTG GTCGAATGGG	3120

CAGGTAGCCG GATCAAGCGT ATGCAGCCGC CGCATTGCAT CAGCCATGAT GGATACTTTC	3180
TCGGCAGGAG CAAGGTGAGA TGACAGGAGA TCCTGCCCG GCACCTCGCC CAATAGCAGC	3240
CAGTCCCTTC CCGCTTCAGT GACAACGTCG AGCACAGCTG CGCAAGGAAC GCCCCTCGTG	3300
GCCAGCCACG ATAGCCGCGC TGCCTCGTCC TGCAGTTCAT TCAGGGCACC GGACAGGTCG	3360
GTCTTGACAA AAAGAACCGG GCGCCCTGC GCTGACAGCC GGAACACGGC GGCATCAGAG	3420
CAGCCGATTG TCTGTTGTGC CCAGTCATAG CCGAATAGCC TCTCCACCCA AGCGGCCGGA	3480
GAACCTGCGT GCAATCCATC TTGTTCAATC ATGCAGAACG ATCCTCATCC TGTCTCTTGA	3540
TCAGATCTTG ATCCCCTGCG CCATCAGATC CTTGGCGGCA AGAAAGCCAT CCAGTTTACT	3600
TTGCAGGGCT TCCAACCTT ACCAGAGGGC GCCCCAGCTG GCAATTCCGG TTCGCTTGCT	3660
GTCCATAAAA CCGCCCAGTC TAGCTATCGC CATGTAAGCC CACTGCAAGC TACCTGCTTT	3720
CTCTTGCGC TTGCGTTTTC CCTTGTCAG ATAGCCCAGT AGCTGACATT CATCCGGGTT	3780
CAGCACCGTT TCTGCGGACT GGCTTCTAC GTGTTCCGCT TCCTTAGCA GCCCTTGCGC	3840
CCTGAGTGCT TGCAGCAGCG TGAAGCTTAA AAAACTGCAA AAAATAGTTT GACTGTGAG	3900
CGGATAACAA TTAAGATGTA CCCAATTGTG AGCGGATAAC AATTCACAC ATTAAAGAGG	3960
AGAAATTACA TATG	3974

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 112 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTTAAAA AACTGCAAAA AATAGTTGA CTTGTGAGCG GATAACAATT AAGATGTACC	60
CAATTGTGAG CGGATAACAA TTTCACACAT TAAAGAGGAG AAATTACATA TG	112